

CLAIMS

What is claimed is:

1. A method of visualizing in two or three dimensions the relationships in diversity space of compounds which are characterized by high-dimensional descriptors comprising the

5 steps of:

- a. selecting a representative subset of the compounds;
- b. calculating or retrieving from a data base the descriptor characteristics for each compound of the subset selected;
- c. generating a distance matrix between all compounds or utilizing a function to
10 generate the distance matrix elements as needed;
- d. computing a hierarchy of clusters, defining cluster centers and partitioning each set of compounds at each level of clustering;
- e. performing a PCA projection onto the first two components;
- f. using a modified stress function which reflects a horizon, running an NLM
15 refinement from the initial PCA coordinates; and
- g. graphically displaying the coordinates of each compound as determined by the NLM refinement.